Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2008 Proceedings

European Conference on Information Systems (ECIS)

2008

Improving the Quality of Article Master Data: Specification of an Integrated Master Data Platform for Promotions in Retail

Axel Winkelman
ERCIS, axel.winkelmann@ercis.uni-muenster.de

Daniel Beverungen
University Muenster, daniel.beverungen@ercis.uni-muenster.de

Christian Janiesch
SAP Research, c.janiesch@sap.com

Joerg Becker
ERCIS, becker@ercis.uni-muenster.de

IMPROVING THE QUALITY OF ARTICLE MASTER DATA – SPECIFICATION OF AN INTEGRATED MASTER DATA PLATFORM FOR PROMOTIONS IN RETAIL

Winkelmann, Axel, University of Muenster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Muenster, Germany, axel.winkelmann@ercis.uni-muenster.de

Beverungen, Daniel, University of Muenster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Muenster, Germany, daniel.beverungen@ercis.uni-muenster.de

Janiesch, Christian, SAP Australia Pty, Ltd, SAP Research CEC Brisbane, Level 12/ 133 Mary Street, Brisbane QLD 4000, Australia, c.janiesch@sap.com

Becker, Joerg, University of Muenster, European Research Center for Information Systems (ERCIS), Leonardo-Campus 3, 48149 Muenster, Germany, becker@ercis.uni-muenster.de

Abstract

Today, meaningful data has to be provided intra-organisationally but also inter-organisationally. One of the major problems of information and data exchange in retail is a lack of structure in the data provided by industry and retailers as well as very heterogeneous and, thus, often poor data quality. Not only master data management but also the automation of promotion processes at the point of sale, e.g. with coupons, can greatly benefit from data which can be analysed automatically and maintained in a collaborative effort. We propose a hierarchical concept for handling article master data by analysing the demands on this data with respect to promotion campaigns. Based on that, we present the architecture and implementation of a promotion platform which utilises this concept to structure articles in retail and to support promotion campaigns more consistently and flexibly than traditional approaches can.

Keywords: Article Master Data, Retail, Couponing, Promotion, Data Quality

1 MASTER DATA QUALITY IN RETAIL

As early as 1992, Tietz pointed out the importance of providing information on-time and at the correct place in retail by bringing up the slogan ‘information beats products’ (Tietz, 1992). Today, meaningful data has to be provided intra-organisationally but also inter-organisationally. On the one hand, retailers rely on article number data coming from manufacturers. On the other hand, manufacturing companies heavily rely on data provided by retailers, e.g. to optimise the flow of goods in their own supply chain with respect to the actual customer demand. Efficient Consumer Response (ECR) as well as Collaborative Planning, Forecasting, and Replenishment (CPFR) are only two examples for recent attempts to provide an integrated data management across sector boundaries. Standardised data formats for transferring data like Electronic Data Interchange (EDI) or centralised master data pools like SINFOS are widely accepted as industry standards by many retailers.

Even so, one of the major problems concerning information and data is a lack of structure in the data exchanged by industry and retail, as well as very variable (and, thus, often poor) data quality. The problem of poor data quality especially affects article master data provided by manufacturing companies, as standardised guidelines for identifiers and abbreviations are missing. As a consequence, identical articles are often handled inconsistently within databases, and different identifiers and multiple abbreviations are used. Thus, the introduction of hierarchical data structures beyond a very generic classification of goods is currently not possible. Due to this lack of structure and quality, article master data can seldom be used in business processes with a high degree of automation, but require the involvement of humans to interpret, update, and restructure data manually.

Based on the semiotic levels of data, Price and Shanks (2005) propose syntactic, semantic, and pragmatic quality criteria. To match the quality of article master data required by retail companies with the functionality implemented within our concept, we conducted a brief survey to ascertain the importance of several quality criteria as perceived by practitioners. We used generic data quality criteria as proposed by Price and Shanks (2005) for designing the questionnaire and adapted them to fit the attributes of article master data. Mainly, we asked two questions. On the one hand, participants were asked to rate the importance of the quality criteria from a retailers’ perspective in general on a five-point Likert scale. On the other hand, participants were asked to appraise the attainment of the quality criteria in their particular company on a five-point Likert scale. Figure 1 shows the results. It has to be noted, though, that the survey was aimed at gaining some insight into practitioners’ perception on data quality of article master data rather than to provide broad empirical support.

Our research addresses the main findings of this survey. In order to evaluate and make the concept usable in a practical context, we designed and implemented a promotional master data platform to support the cooperation between industry and retail. Design science research as a qualitative research method was chosen to guide the research process and concept implementation and evaluation. As Hevner et al. state: “Design science […] creates and evaluates IT artifacts intended to solve identified organizational problems” (Hevner et al., 2004). Thus, this research method is applicable to our research. Furthermore, action research has certain ties to our research endeavor as one of the important aspects of action research is that the researcher usually faces a dichotomy: problem solving interest and research interest (Avison et al., 1999).

The remainder of this paper presents a new hierarchical concept for handling article master data, by analysing the demands on article master data with respect to promotion campaigns in retail. We propose a concept to improve article master data quality. It is applied in the context of automatic coupon processing. Consecutively, we present the implementation of a promotion platform which utilises this concept to conduct promotion campaigns in retail more consistently and flexibly. This platform was conceptualised and implemented as the result of a research project in the area of coupon clearing.

	Criteria proposed by Price/Shanks (2005)	Definition by Price/Shanks (2005)	Derived question for our empirical survey	Importance for retail in general / standard deviation	Implementation in own company / standard deviation
Syntactic	Conforming to meta data	Data follows specified data integrity rules.	The technical integrity of article master data **shall be / is** necessary.	1.69 / 0.8	2.14 / 0.85
Semantic	Mapped completely	Every external phenomenon is represented.	Every listed article **shall be / is** represented by an article master data record.	1.61 / 0.8	1.59 / 0.8
	Mapped unambiguously	Each identifiable data unit represents at most one specifific external phenomenon.	Article master data **shall / does** unambiguously correspond to real articles.	1.35 / 0.6	1.46 / 0.6
	Phenomena mapped correctly	Each identifiable data unit maps to the correct external phenomenon.	Article master data **shall be / is** mapped to the correct manufacturer.	1.81 / 0.9	1.77 / 1.02
	Properties mapped correctly	Non-identifying (i.e. non-key) attribute values in an identifiable data unit match the property values for the represented external phenomenon.	Attributes in article master data (e.g. weight, height) **shall / do** correspond to the attributes of real articles.	1.46 / 0.7	2.41 / 1.05
	Mapped consistently	Each external phenomenon is either represented by at most one identifiable data unit or by multiple but consistent identifyable units whose inconsistencies are resolved within an acceptable time frame.	Article master data **shall be / is** semantically consistent (e.g. use compatible identifiers, no homonyms).	1.61 / 1.00	2.14 / 0.89
	Mapped meaningfully	Each identifiable data unit represents at least one specific external phenomenon.	Article master data **shall be / is** deleted in case the article will no longer be sold.	3.34 / 1.1	3.63 / 1.22
Pragmatic	Accessible (easy, quick)	Data is easy and quick to retrieve.	Article master data **shall be / is** easily retrievable.	1.6 / 1.00	1.81 / 0.79
	Suitably presented	Data is presented in a manner appropriate for their use, with respect to format, precision, units, and the type of information displayed.	Article master data **shall be / is** appropriately presented for their use.	2.27 / 1.00	2.05 / 0.70
	Flexibly presented	Data can be easily manipulated and the presentation customised as needed, with respect to aggregating data and changing the data format, precision, units, or type of information displayed.	Article master data **shall be / is** recombineable according to arbitrary criteria (e.g. for addressing all Coca-Cola products without caffeine in a 0.5 litre bottle).	2.07 / 0.9	2.32 / 1.04
	Timely	The currency (age) of data is appropriate to their use.	Article master data **shall be / is** up-to-date.	1.38 / 0.60	1.73 / 0.82
	Understandable	Data is presented in an intelligable manner.	Article master data **shall be / is** presented understandably.	1.85 / 1.00	2.00 / 0.82
	Secure	Data is appropriately protected from damage or abuse (including unauthorized access, use or distribution).	Article master data **shall only be / is only** accessible (read or write access) to authorized personnel.	2.31 / 0.90	2.14 / 1.08
	Allowing access to relevant meta data	Appropriate meta data is available to define, constrain, and document data.	Article master data **shall be / is** enhanced by meta data (e.g. data-source, validity period).	2.12 / 0.80	2.68 / 1.00

Figure 1. Adaptation of Syntactic, Semantic, and Pragmatic Data Quality Criteria as Proposed by Price and Shanks (2005) to Article Master Data in Retail (n=26).

The paper is organised as follows: In Section 2, we present the status-quo of couponing, especially characteristics of electronic couponing campaigns in retail. We address insufficiencies of traditional coupon concepts and present an innovative concept for master data driven automatic coupon concepts. Due to the innovative character of this concept, tool support is lacking so far. In Section 3, we present our concept of a hierarchical structure of article master data which is incorporated in the design of a promotion platform. Furthermore, the software architecture of the platform is addressed. In Section 4 the functionality of the platform is matched with data quality requirements for couponing campaigns and a summary and outlook is provided.

2 FROM TRADITIONAL TO INTEGRATED COUPONING CONCEPTS

2.1 Traditional Couponing Concepts

Most retailers have used various forms of discount pricing in order to increase customer traffic for multiple reasons: the introduction of a new branch or new products, the clearance sale of excess inventories, the cross-selling of products etc. Thus, couponing is a popular discount instrument. It allows either a direct or indirect discount to be obtained if the redemption conditions of the voucher are met.

The management and processing of basic article data (price, weight, data concerning storage requirements, and packaging) is a crucial prerequisite for the IT-based manufacturing and compilation of products and couponing promotions. This data is referred to as article master data. Although manufacturers often provide article master data electronically from their ERP systems, a sufficiently designed and managed central data source does not exist (cf. Becker & Winkelmann, 2006). Furthermore, it is widely acknowledged that article master data in general suffers from a poor data quality (Kuipers, 2004).

Throughout the 1970s and 1980s, no electronic article master data was needed for conducting couponing campaigns, because coupons were redeemed manually during check-out. However, manual coupon clearing is subject to failure and fraud, and can only be conducted inefficiently by human personnel. In contrast to this, semi-automatic clearing is based on printing an EAN13 barcode on each coupon which identifies the kind and face-value of the coupon.

EAN is a unique article identification number which can be used to unambiguously identify an article or coupon. In retail, EANs are widely used when scanning products and to automatically collect and process sales data. In Germany, EANs starting with the prefix 981 or 982 indicate a coupon (cf. Figure 2). Based on this identification number, the point of sale (POS) system can determine the issuer of the coupon, the face-value of the coupon, and the promotion code. Even so, semi-automatic clearing suffers from the fact that redeeming coupons is not essentially related to buying the specified article. Thus, errors when redeeming coupons and fraud cannot be prevented in this way. Different couponing concepts are discussed in more detail in Winkelmann (2006).

1: Coupon Promotion
2: Number of Decimal Places of 5) (=2)
3: Identification Number
4: Individual Code (Promotional Code)
5: Value (Face Value) (=0.50 Euro)
6: Checksum

	Prefix			Issuer				Promotional number		Value			Check-sum
1)	9	8	1	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11} ,	x_{12}	C
2)	9	8	2	x_4	x_5	x_6	x_7	x_8	x_9	x_{10} ,	x_{11}	x_{12}	C
3)	9	8	1/2	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	C

1), 2) Semi-automatic Coupon Clearing

3) Automatic Coupon Clearing

Figure 2. A Comparison of Different Couponing Concept, with Respect to the EAN Barcode.

2.2 Integrated Coupon Clearing based on Master Data

Automatic coupon clearing becomes ever more important since it avoids the disadvantages of semi-automatic clearing and enables an automated processing without any need to intervene (e.g., at self-checkout counters). The barcode, which is printed on the coupon, consists of a manufacturer ID and a promotion ID (cf. Figure 2, case 3) without any value information. Other information related to the campaign (validity, redemption conditions, rebate etc.) have to be stored in the cash desk (cf. Figure 3). Fraud is prevented, because the clearing is only possible automatically in compliance with the coupon conditions which are stored in the cash desk. Due to the automatic processing it is even possible to issue coupons with complex redemption conditions without imposing additional burden on the cash desk staff.

The preparation of master data and, hence, the consolidation and structuring in product families is a necessary prerequisite for the automated processing of couponing promotions. An improved structuring of promotion master data leads to a better formalization of promotion processes. Within the promotion process, the clearing house is responsible for the consolidation and management of promotional management data from various master data sources (cf. Figure 4).

The concept of EANs and master data management between industry and retail leads to various problems which arise during the collection, processing, and consolidation of promotions. The 13 digit concept of an EAN 13 is used as follows: In general, every producer receives a 7-digit basis number which has to be complemented with another 5-digit internal product code. The 13th digit is a checksum. Each product can be assigned to various EAN 13-codes because each place of production, each product variant, each packaging size etc. has its own EAN code. Therefore, the number of EAN codes for one product can be very high. Although a producer might specify an internal hierarchy concept for his 5 digit internal product code within the EAN 13, the hierarchy is not interpretable for

external partners such as retailers. In general, the EAN coding is a *flat* concept without any taxonomic structure per se.

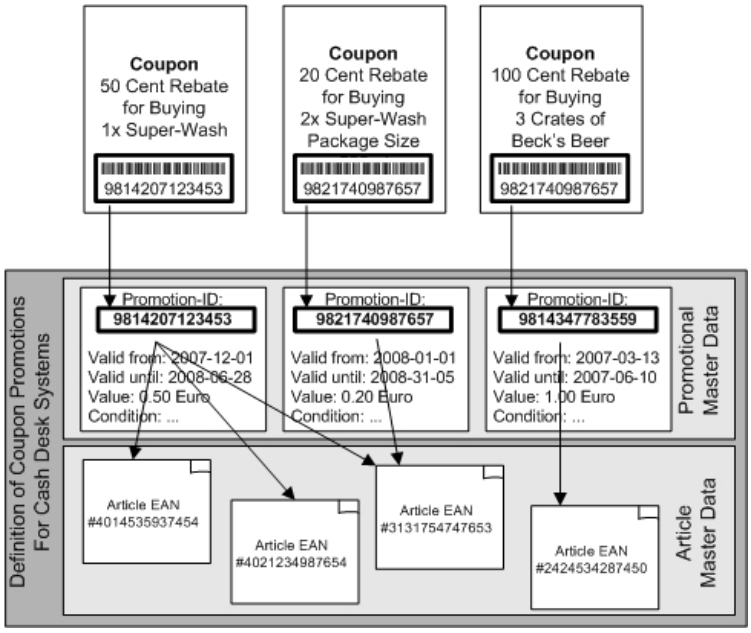

Figure 3. The Roles of Coupon, Barcode, and Master Data in Automatic Coupon Clearing.

As a consequence of the non-existence of an EAN hierarchy, retail and industry have agreed in using a standard product group hierarchy in master data pools. Each article belongs to a product group. Unfortunately, a more detailed structuring within a brand family (e.g., a classification of sub-brands such as *0.5 l Coca-Cola Zero* cans) is not provided. This detailed classification is especially necessary for efficient couponing promotions in order to promote isolated sub-brands such as specific product sizes. Furthermore, for promotions, it is necessary to store article master data for a long time. Some products are available in store for many years. Therefore, a promotion has to include older EANs as well. Additionally, new product EANs might have to be added during the campaign as promotions might last a long period and retailers might have to include new EANs of the promoted products.

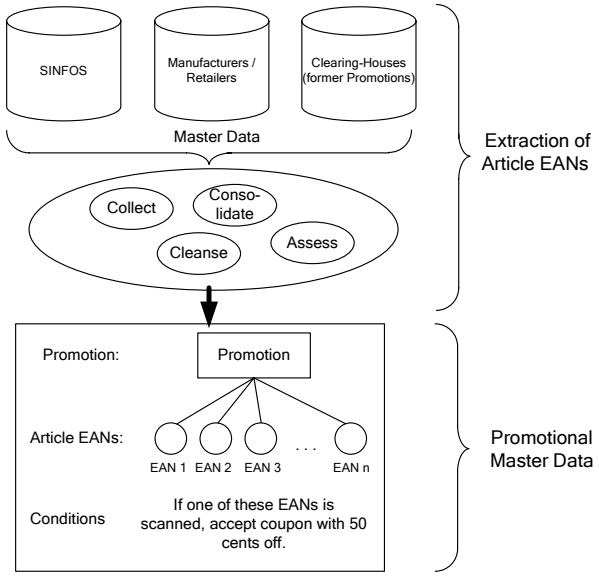

Figure 4. Data Retrieval from Master Data Pools and Data Processing.

To ensure that cash desks are able to use promotion master data, they have to be prepared for each individual cash desk format. Usually, data which is to be used for promotion campaigns is compiled by the clearing house, processed according to the requirements of cash-point systems, and made available to participating retailers.

Consistently managed article master data provided by a central promotion platform can facilitate multilateral cooperation of manufacturers, clearing houses and retailers to design and run promotion campaigns. The cost reduction potential for business processes is expected to be significant and may account for 40 % to 80 % of total costs caused by the inefficiencies of bilateral coordination (Pretzel, 2004). Manufacturers can enter promotion master data into the system (e.g., duration of the campaign, involved articles, and clearing conditions). Retailers can participate in planning couponing campaigns and specify further parameters (e.g., which stores to use for distributing coupons). Thus, the process of planning and issuing a campaign is conducted in a cooperative manner, based on a consistent database.

3 A PROMOTION PLATFORM FOR PROCESSING COUPONS

3.1 Demands towards the Organisation of Master Data on an Integrated Coupon Platform

Conducting integrated promotions such as couponing promotions efficiently, article master data and coupon master data must be managed consistently. Manufacturers and retailers each introduce specific data which has to be thoughtfully organised and provided. One way to achieve this is to design an integrated promotion platform which can be used as a central data repository for manufacturing companies and retailers. Thus, cooperation costs can be reduced and data quality for planning, supervising, updating, and evaluating couponing campaigns can be improved. From a technical point of view, the promotion platform has to be integrated into the current IT landscape of manufacturing companies and retailers, such as POS systems, inventory control systems, or enterprise resource planning (ERP) systems.

As current sources of article master data often suffer from poor data quality and do not provide data in a structured fashion, a reorganisation of data is necessary. On the one hand, article master data must be accessible beyond the currently common four-year data lifecycle. On the other hand, electronic promotion campaigns as a complex promotion concept require the handling of product subsets which can be easily recombined to form abstract product groups. For example, it should be possible to design a campaign to promote all caffeine-free and sugar-free Coca-Cola six-packs with a capacity of 0.5 litres. To allow for an electronic cash-in of the coupons at the point of sale, POS systems need to access the article master data of *all* 0.5 litre six-packs of Coca-Cola which match the specified cash-in conditions. With respect to traditional data sources such as article master data pools, this is currently not possible.

Therefore, one critical prerequisite for conducting electronic promotion campaigns is to allow for a dynamic combination of article master data. To enable employees to set up promotions for dynamically defined groups of articles, article master data first has to be made available in a structured form. Organizing data in a hierarchical order, so called families, is one suitable attempt to provide for this. Hierarchies are taxonomical structures which connect all entities by the principles of subordination (Guarino & Welty 2000).

Therefore, a flexible definition of valid article subsets for couponing campaigns is the main objective for developing an integrated promotion platform. It has to be possible to efficiently compile the articles to be promoted when planning, updating, or evaluating couponing campaigns. Sets of articles must be stored on the promotion platform and have to be assigned to specific couponing campaigns; coupons will then be valid for all articles assigned to the selected groups. To allow for an automatic clearing of coupons, clearing information also has to be stored in the POS system.

3.2 Architecture of an Integrated Promotion Platform

The UML component diagram (Object Management Group Inc., 2007) in Figure 5 displays the software architecture of such a promotion platform. The architecture has been designed according to the principles of service-oriented software systems (Krafzig et al., 2005). Each component exposes a defined interface.

Service-oriented architectures are based on services (Erl, 2005, Papazoglou, 2003). Each of these services contains a reusable application, data, or process logic and encapsulates this functionality by providing a standardised interface. The functionality of the service can be accessed by other applications or services via this interface (Gallas & Schönherr, 2006). Our implementation adheres to the principle of using Service-oriented Architecture (SOA) in the sense of a middleware approach (Schemm et al., 2006) to facilitate inter-organisational cooperation of manufacturers, retailers, and clearing houses.

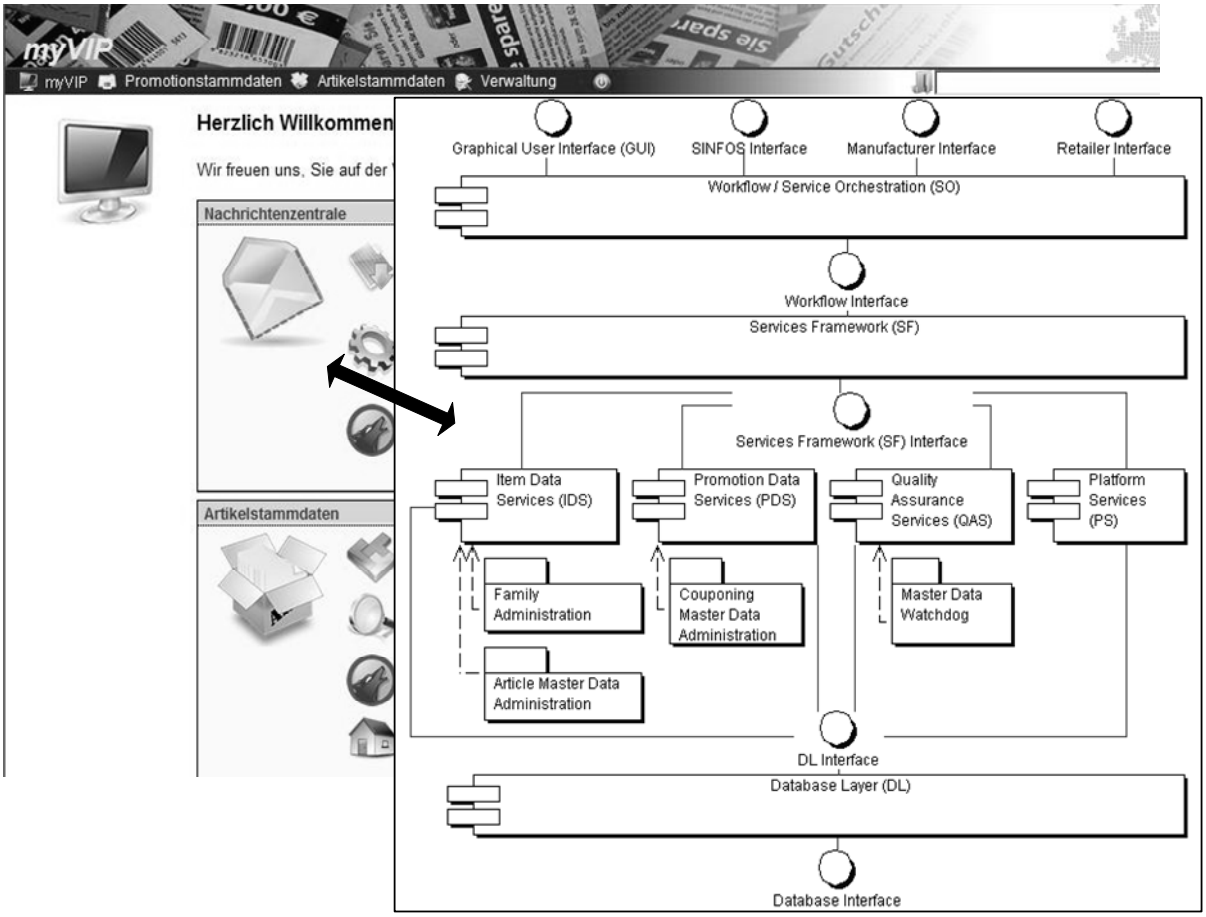

Figure 5. UML Component Diagram of the Services Provided on the Promotion Platform.

The services provided on the promotion platform are managed by a Service Framework (SF) which provides a basis for the integration of these services. As presented earlier, the handling of article master data is crucial for designing and executing promotion campaigns in retail with automatic coupon clearing. Therefore, the platform has to provide adequate functionality to obtain, manage, save, archive, and provide article master data. While data is stored in a relational database (implemented with mySQL), all other functionality is implemented in the Item Data Service (IDS) Layer. Furthermore, IDS provides full-text search functionality for retrieving article master data. The

hierarchical structuring of article master data in abstract sets is the key feature of the platform: It is further detailed in Section 3.3.

Besides article master data, the promotion master data of the promotion campaign (e.g., period, involved retailers, clearing conditions etc.) have to be administered on the promotion platform as well. The platform provides this functionality within the Promotion Data Service (PDS) layer. Functionality includes selecting article master data for promotions (*what* articles to promote), validity period of the promotion (*when* to promote articles), selected retailers (*where* to promote articles), and clearing conditions (*in which* case to accept the coupon during check-out). This data can be accessed by authorised stakeholders, such as manufacturers of the articles involved, or retailers. Cooperative planning and updating of the campaign is provided.

As stated earlier, poor (article master) data quality can seriously impede the success of promotion campaigns. To provide article master data in sufficient quality, the Quality Assurance Service (QAS) layer implements functionality to keep the data repository up-to-date. The layer identifies new article master data in data pools (such as SINFOS) automatically and imports it. Similarity algorithms (based on the built-in full-text search functionality of mySQL 5.0 and the article groupings) help matching new master data records with already existing ones. Thus, it is easier to assess, which coupon campaigns new article have to be added to. This process is very tedious and costly when conducted by users without technical means. It has to be noted though, that the promotion platform presents some hints on updating promotions, but the decision has still to be made by the user. By making use of this functionality, dynamically updating coupon campaigns becomes possible.

The Platform Services (PS) layer provides functionality to securely administrate and provide the content stored on the platform. Access is provided to authenticated users only and restricted to the appropriate data sets. Thus, the secure administration of article master data and promotion master data is assured.

3.3 Hierarchical Organisation of Article Master Data

A structured administration of article master data is key to promote articles in couponing campaigns. This promotion platform organises article master data hierarchically by the IDS layer. Due to the efficiency in searching and updating records in large databases, the hierarchy is conceptualised as a search tree, as shown in the Entity-Relationship diagram in Figure 6.

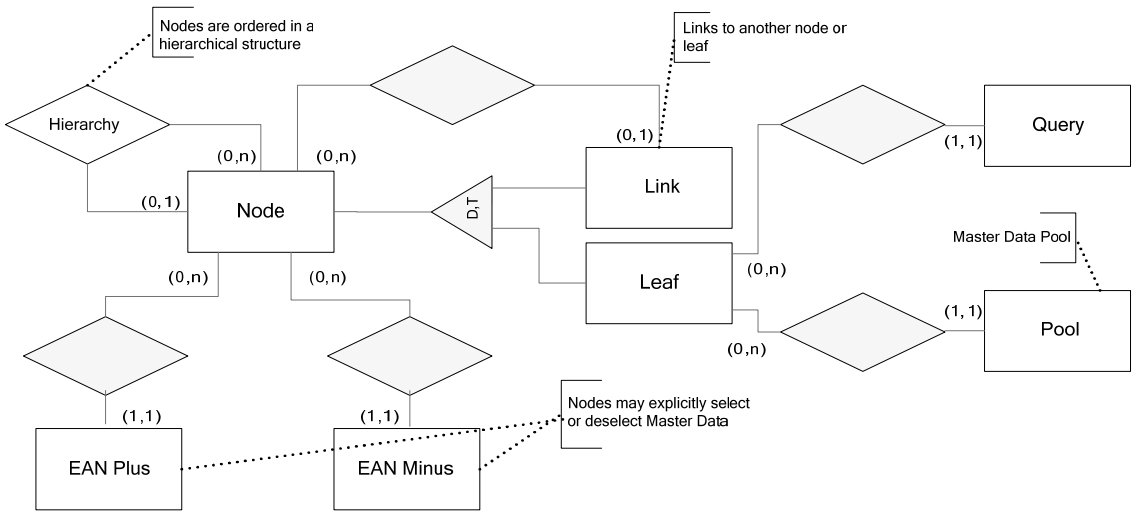

Figure 6. Conceptual Design of a Hierarchical Structure for Article Master Data.

The central elements of the hierarchy are *nodes* which represent abstract groups of articles. The hierarchy itself is conceptualised by the relation *hierarchy* which can be used to rank nodes. The entities *EAN plus* and *EAN minus* represent explicitly included or excluded EANs of articles. Thus, articles stored in the database can be selected or deselected for inclusion in an abstract article group. The entities *leaf* and *link* are special nodes. Links can be used to link from one node to another node. While including links leads to adding all articles of the target node to the source node, excluding links remove all articles from target nodes in the source node. Thus, nodes of erroneous EANs can be specified and then excluded from other nodes in a way that the articles will no longer be part of the promotion campaign. Leafs do not have any subordinate node, but specify *queries* in SQL in the article database. Each query leads to a subset of articles as retrieved from the database. A *pool* of EANs represents all the articles included by queries, while excluded articles are removed from the pool.

This hierarchical article master data structure is considered more useful to conduct promotion campaigns than traditional article master data pools (e.g., SINFOS), because article master data can be combined to form arbitrary and abstract groups, so called families. Therefore, articles to be promoted can be efficiently selected by reusing existing groups of articles or by explicitly deselecting groups of articles that are not to be promoted.

The family concept as implemented in the IDS layer is presented in Figure 7. Each node represents a group of articles (family). A link may point at other nodes. For example, the node ‘Detergent’ contains two including links (one points at the node ‘Ariel’, one at ‘Persil’, both of which are detergent brands). It has to be noted that the general existence of links implies that the data structure is no longer a strict search tree, since several paths from root to particular nodes may be possible. Instead, it can be referred to as an article structure. An excluding link points to faulty EANs which represent articles not to be promoted in the campaign.

At the lowest level of the structure, SQL fragments can be included which represent articles selected from the database according to specified queries. Queries can be used to select articles by name, size, manufacturer, or other criteria from the article database. Furthermore, each node may contain additional valid or invalid articles. This is visualised by grey-shaded rectangles in Figure 7.

In case a family is selected to be promoted, all subordinate nodes will be evaluated and results condensed towards the root element of the article structure. Hence, sets are dynamically processed so that articles can be easily inserted or removed from campaigns. In the example, the node ‘Procter & Gamble’ also includes all EANs represented by the Nodes ‘Ariel’, ‘Compact’, ‘Standard’ and ‘Spring flavour’.

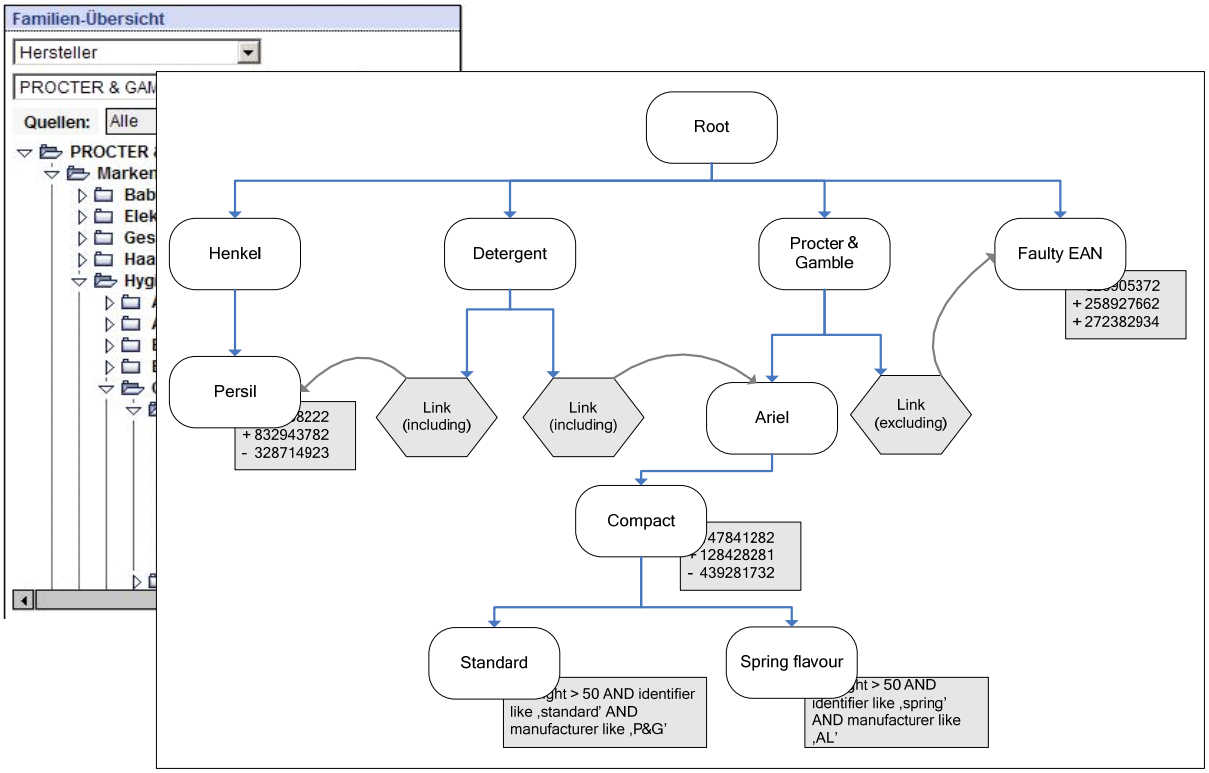

Figure 7. Hierarchical Family Concept of Article Master Data.

4 CONCLUSION: ADDRESSING SYNTACTIC, SEMANTIC, AND PRAGMATIC DATA QUALITY CRITERIA

The presented promotion platform for administrating article master data enables retailer, clearing house and manufacturer to cooperatively manage article master data and coupon master data. The underlying database represents a consistent view on all planned and running promotion campaigns and on the underlying master data at any time.

Key to this concept is a systematic organisation of the underlying article master data. In contrast to this, current data-sources such as SINFOS are vulnerable to inconsistencies and suffer from poor data quality. They cannot provide master data in a structured form that is more detailed than the standard product grouping. This significantly impedes the compilation of article subsets which are needed by promotion campaigns and may finally lead to the failure of the campaign.

Organising article master data in abstract hierarchy sets can better address the needs of automatically conducted promotion campaigns. In our concept, articles are organised in groups (so called families), which can be used as building blocks for other groups. Therefore, the articles to be promoted can be selected more efficiently and kept up-to-date more easily. Furthermore, groups of faulty articles can be explicitly excluded from promotion campaigns by using negative links.

With respect to the syntactic, semantic, and pragmatic data quality criteria adapted from Price and Shanks (2005) in Figure 1, the platform mainly addresses quality criteria related to the pragmatic quality level (cf. Figure 8). Master data records are hierarchically organised in abstract article groups and provide a flexible retrieval of article master data for coupon campaigns. Each dataset can be viewed in detail, displaying each attribute of the dataset, including meta data, such as the source of the dataset and its validity period.

	Criteria proposed by Price/Shanks (2005)	Derived question for our empirical survey	Corresponding functionality of our master data platform
Syntactic	Conforming to meta data	The technical integrity of article master data **shall be / is** provided.	Syntactical data integrity is provided by the Database Management System (DBMS).
Semantic	Mapped completely	Every listed article **shall be / is** represented by an article master data record.	Automatic identification of new master data from several external data sources.
	Mapped unambiguously	Article master data **shall / does** unambiguously correspond to real articles.	[no explicit check possible, as data is obtained from manufacturers' data repositories, such as ERP systems]
	Phenomena mapped correctly	Article master data **shall be / Is** mapped to the correct manufacturer.	[no explicit check possible, as data is obtained from manufacturers' data repositories, such as ERP systems]
	Properties mapped correctly	Attributes in article master data (e.g. weight, height) **shall / do** correspond to the attributes of real articles.	[no explicit check possible, as data is obtained from manufacturers' data repositories, such as ERP systems]
	Mapped consistently	Article master data **shall be / is** semantically consistent (e.g. use compatible identifiers, no homonyms).	Hierarchical administration of master data allows for an improved handling of redundant master data records.
	Mapped meaningfully	Article master data **shall be / is** deleted in case the article will no longer be sold.	Negative links allow for excluding master data records from promotions; no deletion of master data records.
Pragmatic	Accessible (easy, quick)	Article master data **shall be / is** easily retrievable.	Master data can be managed by dynamically selecting and updating article groups for coupon campaigns.
	Suitably presented	Article master data **shall be / is** appropriately presented for their use.	Master data can be flexibly selected for new or modified coupon campaigns.
	Flexibly presented	Article master data **shall be / is** recombineable according to arbitrary criteria (e.g. for addressing all Coca-Cola products without caffeine in a 0.5 litre bottle).	Master data can be managed flexible due to a hierarchical structure.
	Timely	Article master data **shall be / is** up-to-date.	Automatic identification of new master data sets from several external data sources, and computation of similarity indices.
	Understandable	Article master data **shall be / is** presented understandably.	Each master data set can be viewed in detail.
	Secure	Article master data **shall only be / is only** accessible (read or write access) to authorised personnel.	Role-based user access control: Only authorised users may view data.
	Allowing access to relevant meta data	Article master data **shall be / is** enhanced by meta data (e.g. data-source, validity period).	Information on the master data source, computed similarity of master data entries.

Figure 8. Functionality of the Promotion Platform to Address Syntactic, Semantic, and Pragmatic Quality Criteria, as Proposed by Price and Shanks (2005).

Thus, the concepts to enhance the quality of article master data, as presented in this paper, focus on quality dimensions of the conceptual view (Redman, 1996). As opposed to this focus, semantic quality categories (i.e. quality dimensions of data values, as focused by Cappiello et al., 2003) cannot be addressed by the platform in depth, because article master data is imported from external data sources (i.e. manufacturers' ERP systems). Therefore, manufacturers own the data and are responsible to guarantee its accuracy, completeness, currency, and value consistency. As a consequence, incomplete data is hard to be completed by others ex post. Nevertheless, in our platform new data sets are identified automatically and imported from external data sources, to keep master data entries up-to-date. Erroneous or incomplete data can be excluded from coupon campaigns by using negative links.

On a syntactical level, data consistency (Weikum & Vossen, 2002) is provided by the database management system. While we used a mySQL implementation, other databases can easily be used as the implementation provides a non-proprietary database interface.

So far, we were not able to identify competing tool-based approaches. In the US, the problem of product and coupon matching is solved with the help of product group numbers ("family codes") within the barcodes. As every product can only be assigned to one product family, the concept is not very flexible and does not allow the creation of sub-categories at all.

We have identified further research potential on interoperability issues between heterogeneous data sources such as other industry-specific master data pools. Also, interfaces to common retail ERP systems need to be made more sophisticated. While the implementation is currently evaluated in cooperation with an industry partner, long term benefits and large scale testing is necessary to quantify the benefits which can currently only be measured qualitatively.

References

Avison, D., Lau, F., Myers, M. and Nielsen, P. A. (1999). Action Research. Communications of the ACM, 42 (1), pp. 94-97.

Becker, J. and Winkelmann, A. (2006). Handelscontrolling: Optimale Informationsversorgung mit Kennzahlen. Springer, Berlin.

Cappiello, C., Francalanci, C. and Pernici, B. (2003). Time-Related Factors of Data Quality in Multichannel Information Systems. Journal of Management Information Systems, 20 (3), pp. 71-91.

Erl, T. (2005). Service-oriented Architecture: Concepts, Technology, and Design. Prentice Hall, Englewood Cliffs, NJ.

Gallas, B. E. and Schönherr, M. (2006). Service Management als Grundlage Service Orientierter Architekturen: Ein SOA-Testszenario. In Unternehmensarchitekturen und Systemintegration. 2nd Edition (Aier, S. and Schönherr, M. Eds.). In Enterprise Architecture, pp. 221-245, Gito, Berlin.

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28 (1), pp. 75-105.

Krafzig, D., Banke, K. and Slama, D. (2005). Enterprise SOA: Service-oriented Architecture Best Practices. Prentice Hall, Upper Saddle River, NJ.

Kuipers, P. (2004). Internal Data Alignment: Data in Dire Need of a Spring-clean. Elsevier Food International, 7 (3), pp. 74-79.

Object Management Group Inc. (2007). Unified Modeling Language 2.0. Downloaded from http://www.uml.org/#UML2.0 on 2007-11-26.

Papazoglou, M. P. (2003). Service-Oriented Computing: Concepts, Characteristics and Directions. In Proceedings of the 4th International Conference on Web Information Systems Engineering (WISE), pp. 3-12, Roma.

Pretzel, J. (2004). Nutzenpotenziale in der Logistik: Chancen für die Wertschöpfung, Köln.

Price, R. and Shanks, G. (2005). A Semiotic Information Quality Framework: Development and Comparative Analysis. Journal of Information Technology, 20, pp. 88-102.

Redman, T. C. (1996). Data Quality for the Information Age. Artech House, Boston, MA.

Schemm, J., Heutschi, R., Vogel, T., Wende, K. and Legner, C. (2006). Serviceorientierte Architekturen: Einordnung im Business Engineering, University of St.Gallen.

Tietz, B. (1992). Einzelhandelsperspektiven für die Bundesrepublik Deutschland bis zum Jahre 2010. Vol. 1. Deutscher Fachverlag, Frankfurt.

Weikum, G. and Vossen, G. (2002). Transactional Information Systems: Theory, Algortihms, And The Practice of Concurrency Control and Recovery. Academic Press, London.

Winkelmann, A. (2006). Integrated Couponing: A Process-Based Framework for In-Store Coupon Promotion Handling. Dissertation. University of Logos, Berlin.